Volume: 4 Issue: III Month of publication: March 2016
DOI:

Volume 4 Issue III, March 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Secure Deduplication with Authorized Duplicate Check Using Convergent Encryption in Cloud

V.P.Selvanathan[1], Mrs.M.Suguna[2]
Department of CSE, Kumaraguru College of Technology

Abstract*-- Increased growth of cloud storage turns the focus of cloud service providers to save their resources to earn more profit by renting them to the multiple cloud users. One of the main factors that utilizes more space and causes resource unavailability for new cloud users is the presence of duplicate copies of data files of different owners. Data duplication is one of the most efficient technique which attempts to reduce the multiple redundant copies of single file into the same file. To provide the memory consumption in the considerable manner, in this work convergent key encryption based duplicate reduction is introduced. This mechanism will encrypt the data file by using the part of content of same file which would be stored along with the encrypted copy of data file. A new user, in order to upload a new file needs to send the convergent copy of own file to the cloud service providers which will be matched with the other copy of file. If it already exists, then the redundant storage of files can be avoided considerably. The experimental results of this work prove that the proposed approach leads to a better result than the existing work in terms of accuracy.

***Keywords*—Cloud Computing, Encryption, duplication, convergent key, service providers**

I. INTRODUCTION

Cloud storage offers highly-available virtually infinite and quick to scale storage with its pay as you go model in recent years it has attracted new customers by the score. Coupled with dropping prices the cloud paradigm has turned storage into a commodity. The decreasing cost of storage media the use of multi tenancy competition between cloud providers and the efficient use of the storage backend through compression and deduplication can be listed amongst the reasons for low price high quality cloud services such as cloud storage services. As an illustration multiple copies of popular content need to be stored only once upon the upload subsequent upload requests can be discarded and only require establishing a link from the uploading user to the original copy of the content. Deduplication can be performed very effectively at both file and block- level deduplication for the same application by the same vendor depending on the setup and the input dataset.

II. DEDUPLICATION

Deduplication can take place at the client side or at the server side. If deduplication is triggered at the client side it is more efficient as it saves upload bandwidth. This is especially for service providers due to the fact that network activity is the most energy consuming task for cloud. To keep away from the transmission of the complete content but still allowing to check for its existence at the server side clients are usually asked to generate a much shorter version of and to use that digest to uniquely identify. The standard approach is to interpret the upload of a digest by a client as a proof that the client essentially owns that. In the work patterned the protection weaknesses hidden behind approach. Primary the privacy and reveal gentility of users of a storage system can be compromised by an attacker that checks if another user has already uploaded a by trying to upload it as well. If the upload does not take place it means the server already stores it. This can be extremely dangerous if the is very rare or private. Second deduplication can be abused to turn the service provider into an underground direct. Two join together users with no direct connectivity can establish a protocol to exchange information stealthily. For instance to exchange one bit of information one of the users checks if a previously agreed has been uploaded or not during a certain time window If the was uploaded the user can consider that a 1 has been transmitted. Finally a cloud storage service can be used as a content distribution network (CDN). In such a case a user can share large with other users just by exchanging the consequent assimilate.

III. LITEARTURE REVIEW

A. Improved Proxy Re-Encryption Schemes with Applications to Secure Distributed Storage - Giuseppe Ateniese, Kevin Fu, Matthew Green, Susan Hohenberger

The Cipher text is computed using proxy re-encryption technique that allows a proxy under which Bob can decrypt Alice's public key. Alice might wish to temporarily forward encrypted email to the colleague Bob, without giving a secret key. In this case, Alice the delegator could designate a proxy to re-encrypt a particular incoming mail into concern format that Bob the delegate can decrypt using the own secret key. Alice could simply issue the secret key to the proxy, but this needs an unrealistic level of trustworthy in the proxy. Several efficient proxy re-encryption schemes has been proposed, that offer security improvements over earlier approaches. The predominant advantage of our schemes is that they are unidirectional and thus do not require delegators to confess all of their secret key to anyone—or even to interact with the delegate—in order to allow a proxy to re-encrypt their cipher text. In our schemes, only a limited quantity of trust is fitted in the proxy. This enables many applications that will not be practical if the proxy is needed to be fully trusted. The first actual performance measurements of applications are provided using proxy re-encryption. To demonstrate the practical utility of our proxy re-encryption schemes, an implementation of proxy re-encryption used is measured in a secure file system. This system uses a centralized access control server in order to manage access to encrypted content stored on distributed, untrusted replicas. Proxy re-encryption has been allowed for centrally-managed access control without awarding full decryption rights to the access control server.

B. Fade: Secure Overlay Cloud Storage With File Assured Deletion - Y. Tang, P. P. C. Lee, J. C. S. Lui, And R. Perlman

In this work, it presents FADE, a secure overlay cloud storage system that safeguard file assured deletion and works seamlessly upon today’s cloud storage services. FADE dissociates the management of encrypted data and encryption keys and other files, so that encrypted data remains in untrustworthy cloud storage providers, while encryption keys are independently undertaken by a key manager service, which can be enforced using a quorum scheme to ensure trustworthiness. FADE generalizes file assured deletion which is time-based, into a more fine-grained approach known as policy based file assured deletion, where the files are associated with more flexible file access policies (e.g., read/write permissions, time expiration of authorized users) and are assuredly deleted while the associated file access policies are cancelled and become obsolete.

1) Policy Metadata : The policy metadata accommodate the blueprint of the Boolean combination of policies and alike encrypted cryptographic keys. Here, the assumption that each single policy is stated by a unique 4-byte integer identifier is made. To represent a Boolean combination of policies, it is expressed in disjunctive canonical form, i.e., the parting (OR) of conjunctive policies, and practice the characters ‘*’ and ‘+’ to stand for the AND and OR operators. Then the policy metadata is updated as a separate file to the storage cloud.

C. Privacy-Preserving Public Auditing For Secure Cloud Storage - C. Wang, Q. Wang, K. Ren, And W. Lou

The improper advantages in the IT history are on-demand self-service, everywhere network access, location of data independent resource pooling, usage-based pricing, brisk resource elasticity and transference of risk. As a disruptive technology with intelligent implications, Cloud Computing is transforming the nature of how the businesses use information technology. One fundamental theme of this paradigm shifting is that which data is being centralized or outsourced to the Cloud. To address these problems, this work exploit the approach of public key based homomorphic linear authenticator which facilitate TPA to perform the auditing not with the demanding the local copy of data and thus drastically reduces the communication and estimation overhead as compared to the forthright data auditing approaches. By integrating the HLA with random masking, protocol assures that the TPA could not able to learn any knowledge about the data content stored in the cloud server during the profitable auditing process.

D. Distributed Computing With Load-Managed Active Storage - R. Wickremesinghe, J. Chase, And J. Vitter

ASUs allow processing capacity to scale actually with the storage capacity. To reduce data movement across the interconnection if searching, filtering, or read/modify/write steps execute directly on ASUs is necessary. This allows gathering of a much larger numbers of drives regarding each network port, and it can improve host processing performance since in host memory data movement is usually a main drain on host CPU resources. However, ASUs introduce new distributed computing challenges regarding to scrutinizing the planning of application process to ASUs, coordinating functions across hosts and ASUs, and sharing of ASU resources. This work explores a programming model to support computation on ASUs, with an emphasis on flexible resource management at the system level. Three factors motivate our approach. First, network storage is a shared resource, and storage-based computing should not occur if it interferes with storage access for other applications. Secondly, an asymmetric parallel processing model is represented by the ASUs, the processing power available in the storage hierarchy may vary widely across configurations, and applications should configure to make the best use of the parallelism available. Thirdly, the active storage offers a potential for local control over data movement and access order for the optimization of storage performance; the ordering constraints must expose the application structure precisely, so that ASUs may realign operations when it is very useful to do so. It proposes a model of load-managed active storage, which takes effort in combining computation with storage access in a way which permits the system to presume the backlashes of loading off computation to ASUs so that it may configure the application to equalize hardware capabilities and conditions of load. This approach extends a well-established model for external memory programming. In the extensified version, computations are embedded in a dataflow style by constituting streaming primitive's functions those function on sets or streams of fixed-size records passing through them. One premise of this work is that the techniques in order to formulate external memory algorithms uncover parallelism that the system may exploit by mapping functions to ASUs. A key goal of this approach is to permit load-managed storage that is active. Functions accomplish bounded computation has I/O access as a side effect; the mapping of functions to ASUs and hosts is configurable and probably dynamic. In order to denote the potential of load-managed active storage, the results from a reconfigurable active storage mergesort algorithm are established.

E. Patient Controlled Encryption: Ensuring Privacy Of Electronic Medical Records - Josh Benaloh, Melissa Chase, Eric Horvitz, And Kristin Lauter

Encryption schemes constituting of strong security features will guarantee that the patient's privacy is protected. However, sticking to a simple encryption scheme can impede with the desired functionality of health record systems. Mainly, it like to establish encryption, but also to support such desirable functions as granting users to share partial access rights with some others and also to accomplish various searches over their records. Thus take into account encryption schemes that enable patients to entrust partial decryption rights, and also that allow patients to search over their health data. Further it is to be proposed that a design that it specify to as Patient Controlled Encryption (PCE) as a result to safeguard and private medical records storage of patients. PCE which enables to share records among healthcare providers and doctors. The architecture of the system is based on a hierarchical encryption model. The patient's record is subdivided into a hierarchical structure, each portion of which is encrypted with a selective key. This patient requires storing a root secret key, from which the subkeys is derived. The patient can selectively issue the subkeys for decryption of various portions of the record.

F. Attribute-Based Encryption For Fine-Grained Access Control Of Encrypted Data - Vipul Goyal, Omkant Pandey, Amit Sahai, Brent Waters

The attribute-based encryption (ABE) is anticipated to be an appreciating tool for carryout fine-grained access control in new public key primitive. To further address the interest of user access privacy, privacy-aware ABE schemes are being grown to achieve hidden access policy recently. Now, no ABE scheme can completely forbid the problem of illegal key sharing among users. In this paper, this problem is solved by firstly proposing the notion of accountable, anonymous, and cipher text-policy ABE (CP-A 3 BE, in short) and then giving out a concrete construction. This is improving the state-of-the-art of unidentified CP-ABE to obtain shorter public parameters and cipher text length. In the proposed CP-A 3 BE construction, user accountability can be achieved in black-box model by enclosing additional user-specific information into the attribute private key provided to that user, while still maintaining hidden access policy.

G. Fuzzy Identity Based Encryption - Amit Sahai, Brent Waters

A new type of Identity Based Encryption (IBE) scheme is popularized that called as Fuzzy Identity Based Encryption. A Fuzzy IBE scheme allows for a private key for an existence id to decrypt a occurrence cipher text encrypted with another existence id if and only if the existence id and id$'$ are close to each other as measured by some metric. A Fuzzy IBE scheme can be applied to permit encryption using biometric measurements as existence. The fault-tolerance of a Fuzzy IBE scheme is literally what allows for the use of biometric identities, which constitutionally contain some amount of noise during each measurement. In this paper, a construction of a Fuzzy IBE scheme is proposed that uses groups with efficiently computable bilinear maps. Moreover, the construction does not use Random Oracles. The security of the scheme is proved under the Selective-ID security model.

IV. CONCLUSION

Dekey is a dynamic and steady convergent key management strategy for highly secure deduplication. Dekey applies deduplication amid convergent keys and assigns convergent key shares over multiple key servers, while sustain semantic security of convergent keys and confidentiality of redistributed data. In this paper, an important security concern is addressed in cross-user client-side deduplication of encrypted files in the cloud storage: confidentiality of users' sensitive files against both outside attackers and the honest-but-curious cloud storage server in the bounded leakage model. On technique aspect, the convergent encryption method is enhanced and generalized and the resulting encryption scheme could support client-side deduplication of encrypted file in the bounded leakage model.

REFERENCES

[1] Giuseppe Ateniese, Kevin Fu, Matthew Green, Susan Hohenberger, "Improved Proxy Re-Encryption Schemes with Applications to Secure Distributed Storage", ACM Transactions on Information and System Security (TISSEC), Volume 9 Issue 1, February 2006.

[2] Yang Tang, Patrick P. C. Lee, John C. S. Lui, and Radia Perlman, "FADE: Secure Overlay Cloud Storage with File Assured Deletion", Dependable and

Secure Computing, IEEE Transactions, 2011.

[3] Cong Wang, Sherman S.-M. Chow, Qian Wang, Kui Ren, and Wenjing Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage", Computers, IEEE Transactions, 2013.

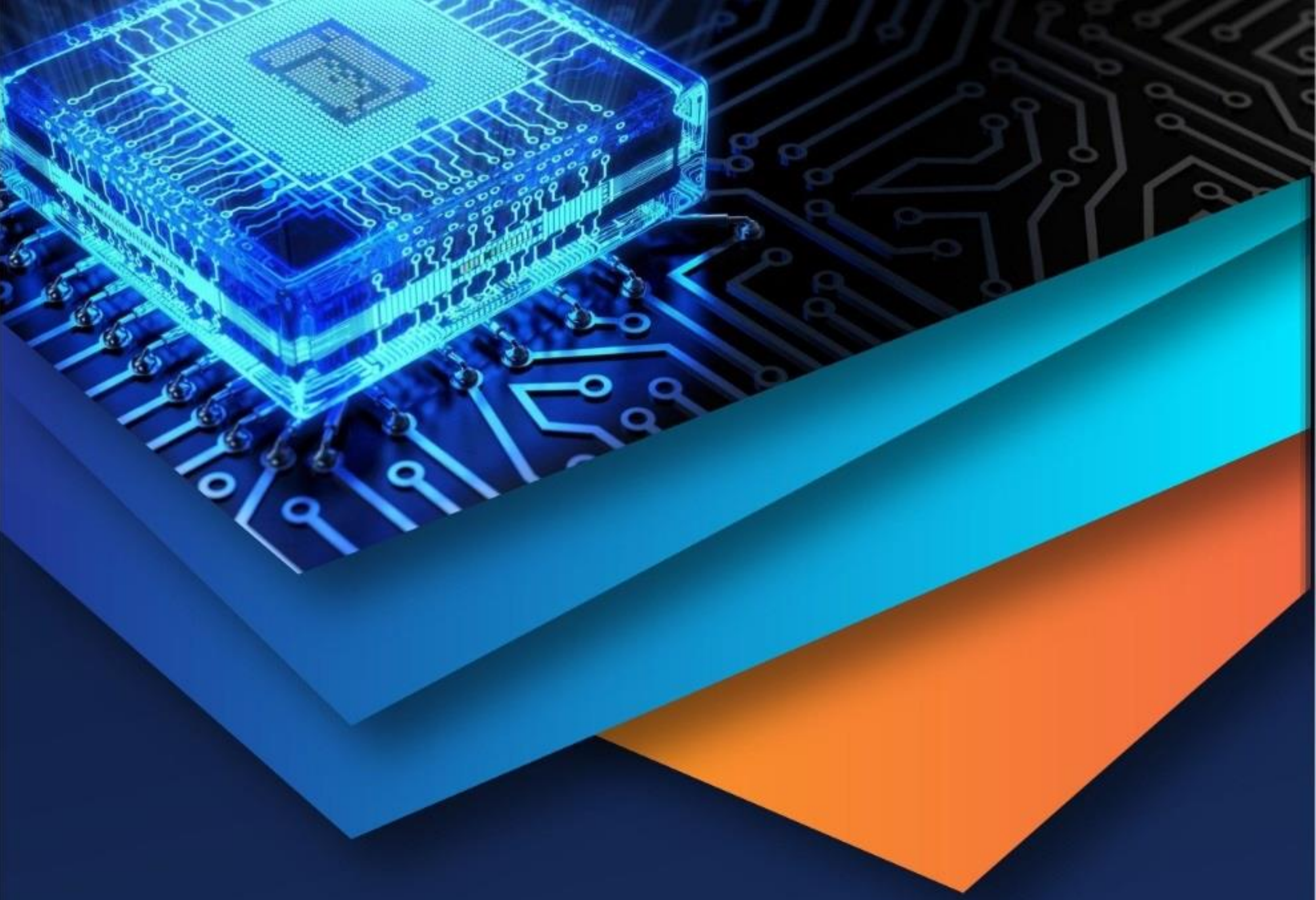
[4] Brent Welch, Marc Unangst, Zainul Abbasi, Garth Gibson, Brian Mueller, Jason Small, Jim Zelenka, Bin Zhou Panasas, Carnegie Mellon, "Scalable Performance of the Panasas Parallel File System", This paper was published in the proceedings of the 6th USENIX Conference on File and Storage Technologies (FAST '08), San Jose, California, February 26-29, 2008.

[5] Rajiv Wickremesinghe, Jeffrey S. Chase, Jeffrey S. Vitter, "Distributed Computing with Load-Managed Active Storage", High Performance Distributed Computing, 2002. HPDC-11 2002. Proceedings. 11th IEEE International Symposium on 2002.

[6] Josh Benaloh, Melissa Chase, Eric Horvitz, and Kristin Lauter, "Patient controlled encryption: ensuring privacy of electronic medical records", Proceedings of the 2009 ACM workshop on Cloud computing security.

[7] Vipul Goyal, Omkant Pandey, Amit Sahai, Brent Waters, "Attribute-Based Encryption for Fine-Grained Access Control of Encrypted Data", Proceedings of the 13th ACM conference on Computer and communications security, 2012.

[8] Amit Sahai, Brent Waters, "Fuzzy Identity-Based Encryption", Volume 3494 of the series Lecture Notes in Computer Science, 2011.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY